

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平8-87297

(43)公開日 平成8年(1996)4月2日

(51)Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L	5/04	F		
	3/00	H		
	5/02	J		

審査請求 未請求 請求項の数12 O L (全 11 頁)

(21)出願番号 特願平6-225396

(22)出願日 平成6年(1994)9月20日

(71)出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中1015番地

(72)発明者 松本 達郎

神奈川県川崎市中原区上小田中1015番地

富士通株式会社内

(74)代理人 弁理士 河野 登夫

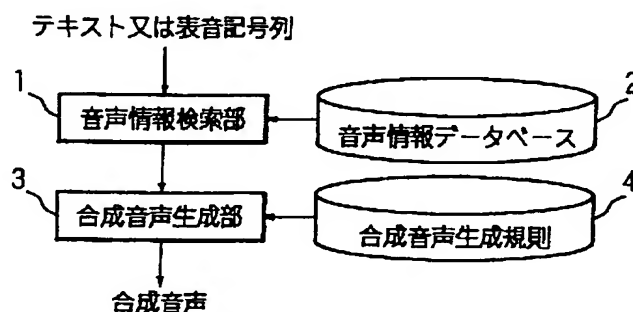
(54)【発明の名称】 音声合成システム

(57)【要約】

【目的】 実音声に近い明瞭度、韻律を持った自然性の高い合成音声を得る。

【構成】 テキスト又は表音記号列が入力されると、音声情報検索部1は実音声进行分析し、抽出した音声特徴量及びこれに対応する発声内容を格納した音声情報データベース2中に、入力テキスト又は入力音声記号列に一致する発声内容が存在するかどうかを検索し、一致する発声内容が存在する場合はこれを合成音声生成部3へ渡す。合成音声生成部3は音声情報に応じた処理を施して合成音声を生成し、また一致する発声内容が存在しない場合は入力テキスト又は入力表音記号列をそのまま合成音声生成部3へ渡し、合成音声生成部3により合成音声生成規則4に基づき合成音声を生成する。

本発明の原理図1



【特許請求の範囲】

【請求項 1】 入力テキスト又は入力表音記号列を合成音声に変換して出力する音声合成システムにおいて、実音声から抽出した音声特徴量を格納した音声情報データベースと、入力テキスト又は入力表音記号列と対応する前記音声情報データベースに格納されている音声特徴量を検索する検索手段と、検索の結果、音声情報データベースに該当する音声特徴量が存在した場合はその音声特徴量に基づいて音声を構成する手段と、該当する音声特徴量が存在しなかった場合は予め定めた規則に従って合成音声を生成する手段とを具備することを特徴とする音声合成システム。

【請求項 2】 音声情報データベースとして音声特徴量中の実音声から抽出した音響的特徴量を格納した音響情報データベースを用いることを特徴とする請求項 1 記載の音声合成システム。

【請求項 3】 音声情報データベースとして、音声特徴量中の実音声から抽出した韻律的な特徴量を格納した韻律情報データベースを用いることを特徴とする請求項 1 記載の音声合成システム。

【請求項 4】 音声情報データベースとして、実音声から抽出した音響的特徴量を格納した音響情報データベース及び実音声から抽出した韻律的な特徴量を格納した韻律情報データベースを用いることを特徴とする請求項 1 記載の音声合成システム。

【請求項 5】 音響情報データベースとして、音声波形を格納した音声波形データベースを用いることを特徴とする請求項 2 又は 4 記載の音声合成システム。

【請求項 6】 音響情報データベースとして、スペクトル、声道断面積又はフォルマント周波数を格納した音声パラメータデータベースを用いることを特徴とする請求項 2 又は 4 記載の音声合成システム。

【請求項 7】 音響情報データベースとして、音声波形データベース及び音声パラメータデータベースを用いることを特徴とする請求項 2 又は 4 記載の音声合成システム。

【請求項 8】 韻律情報データベースとして、音声長、音声強度、基本周波数のうちの、音声長のみ、又はいずれか 2 つ以上を格納したデータベースを用いることを特徴とする請求項 3 又は 4 記載の音声合成システム。

【請求項 9】 韻律情報データベースとして音声長及び音声強度を格納した音声長・音声強度データベースを用いることを特徴とする請求項 3 又は 4 記載の音声合成システム。

【請求項 10】 韻律情報データベースとして、音声長及び基本周波数を格納した音声長・F0 データベースを用いることを特徴とする請求項 3 又は 4 記載の音声合成システム。

【請求項 11】 韻律情報データベースとして音声長のみを格納した音声長データベースを用いることを特徴と

する請求項 3 又は 4 記載の音声合成システム。

【請求項 12】 韻律情報データベースとして、音声長・音声強度・F0 データベース、音声長・音声強度データベース、音声長・F0 データベース又は音声長データベースのいずれか 2 以上のデータベースを用いることを特徴とする請求項 3 又は 4 記載の音声合成システム。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は任意の入力テキスト、又は入力表音記号列を合成音声に変換し、出力する音声合成システムに関する。

【0002】

【従来の技術】 図 10 は従来の音声合成システムの主要構成を示すブロック図であり、任意の入力テキストを合成音声に変換し、出力する場合を示している。図中 304 は言語処理部を示しており、言語処理部 304 はテキスト入力に対し、単語辞書 303 を適用して入力テキスト中における単語の読み、アクセントの位置を決定すると共に、文の構造を解析し、イントネーションを制御するための制御情報を作成し、これを音声長パターン生成部 300 へ渡す。通常単語の読み、アクセント位置、イントネーション制御情報はいずれも表音記号列と呼ばれる記号列で表現される。

【0003】 音声長パターン生成部 300 は、音声長生成規則 400 を適用して表音記号列を各音の長さ情報（音声長）パターンに変換し、この音声長パターン及び表音記号列を F0 パターン生成部 301 へ渡す。F0 パターン生成部 301 は、F0 生成規則 401 を適用して表音記号列と音声長パターンとに基づきアクセント、イントネーションに相当する物理量である基本周波数（F0）の時間変化パターン、所謂 F0 パターンを生成し、表音記号列、音声長パターン及び F0 パターンを音声パラメータパターン生成部 310 へ渡す。

【0004】 音声パラメータパターン生成部 310 は、音声パラメータ生成規則 410 を適用して表音記号列（特に読み情報）、音声長パターン、F0 パターンに基づき音声パラメータパターンを生成し、これを波形生成部 311 へ渡す。なおここに音声パラメータパターンは通常 RARCOR 係数（声道断面積に対応する係数）、又はフォルマント（声道共振）周波数等の音声特徴量と音源信号の時間変化パターンである。また音声波形の短い単位を接続して合成音声を生成する、所謂波形編集方式の場合は、音声波形の短い単位である素片波形の種類、接続タイミング等の接続情報である。

【0005】 波形生成部 311 は、渡された音声パラメータパターンに基づいて実際のデジタル音声波形を生成し、これを DA 変換部 5 へ渡す。例えば音声パラメータパターンが PARCOR 係数である場合には、波形生成部 311 は PARCOR フィルタと音源生成部とで構成され、フィルタを音源信号で駆動する処理を行う。ま

た、波形編集方式の場合は素片波形を適切な位置に配し、これらを滑らかに接続する処理を行う。DA変換部5は波形生成部311で生成されたデジタル音声波形をアナログ音声波形に変換し、合成音声として出力する。

【0006】

【発明が解決しようとする課題】ところでこのような従来の音声合成システムでは、全ての音響的、韻律的処理を予め用意してある各規則に基づいて行っているが、圧縮された状態の音響情報に基づいて合成音声生成するため、合成音声の品質の劣化を免れ得ず、また規則によって音声長パターン、FOパターン及び音声パラメータパターンを生成するため、不自然なアクセント、イントネーションになる等の問題があった。

【0007】本発明の第1の目的は、音響情報データベース、韻律情報データベース等、複数種類の音声情報データベースを用意して入力テキストがこれらデータベースに登録されているテキストである場合には、データベースから引き出した情報に基づいて音声を再構築し、出力することで再生音声の品質向上を図ることにある。

【0008】本発明の他の目的は、音響情報データベースとして音声波形データベース及び／又は音声パラメータデータベースを、また韻律情報データベースとして音声長データベース及び／又は音声長・音声強度データベース及び／又は音声長・FOパターンデータベースを用いることで様々な入力テキスト、又は入力表音記号列に対応した音声合成を可能とすることにある。

【0009】

【課題を解決するための手段】本発明の原理を説明する。図1は本発明の第1の原理を示す原理図1であり、図中1は音声情報検索部、3は合成音声生成部を示している。

【0010】音声情報検索部1は、テキスト又は表音記号列の入力があると人が発した音声である実音声から抽出した各種の音声特徴量、及びこれと対応する発声内容（音声でどのように話されたかを示すラベル）を格納してある音声情報データベース2中に、これら入力テキスト又は入力表音記号列と一致する発声内容が存在するか否かを検索する。

【0011】なお、ここに音声情報とは音声波形、音声パラメータ、FO、音声強度及び音声長等をいう。また音声特徴量とは音声の物理的な特徴量、所謂音声パラメータであり、一般的には音声周波数領域の特徴量を指し、これにはスペクトル（周波数強度）フォルマント（声道共振周波数）、LPC（線形予測係数）、PARCOR係数（声道断面積に対応する係数）等がある。

【0012】検索の結果、一致する発声内容が存在する場合はこれを合成音声生成部3へ渡し、また一致する発声内容が存在しない場合には、入力テキスト又は入力表音記号列をそのまま合成音声生成部3へ渡す。合成音声

生成部3は、合成音声生成規則4に基づいて、入力テキスト又は入力表音記号列から合成音声を生成する。

【0013】第1の発明は、この原理に基づく発明であり、入力テキスト又は入力表音記号列を合成音声に変換して出力する音声合成システムにおいて、実音声から抽出した音声特徴量を格納した音声情報データベースと、入力テキスト又は入力表音記号列と対応する前記音声情報データベースに格納されている音声特徴量を検索する検索手段と、検索の結果、音声情報データベースに該当する音声特徴量が存在した場合はその音声特徴量に基づいて音声を構成する手段と、該当する音声特徴量が存在しなかった場合は予め定めた規則に従って合成音声を生成する手段とを具備することを特徴とする。

【0014】図2は本発明の第2の原理を示す原理図2であり、図中10は音響情報検索部、30は韻律情報生成部、31は音響情報生成部を示している。なお、音響情報とは、音声情報のうちの音声波形の如き時間領域の情報、スペクトルの如き周波数領域の情報を意味し、また音響的特徴量とはこれら情報のうちの音声合成上の有意な情報を意味する。また、韻律情報とは、音声情報のうちの韻律（イントネーション、アクセント、リズム、強度）についての情報を意味し、また韻律的特徴量とはイントネーション、アクセントに対応する韻律の物理的特徴量である基本周波数（FO）、リズムに対応する音声長、強度に対応する音声強度をいう。

【0015】音響情報検索部10はテキスト又は表音記号列が入力されると実音声から抽出した様々な音響的特徴量及びこれと対応する発声内容が格納されている音響情報データベース20中に入力テキスト又は入力表音記号列と一致する発声内容が存在するか否かを検索する。一致する発声内容が存在する場合は、これを直接音響情報として出力し、また一致する発声内容が存在しない場合には入力テキスト又は入力表音記号列をそのまま韻律情報生成部30へ渡す。

【0016】韻律情報生成部30は、韻律情報生成規則40に基づいてテキスト又は表音記号列から韻律情報を生成し、これを音響情報生成部31へ渡す。音響情報生成部31は音響情報生成規則41に基づいて韻律情報から音響情報を生成する。

【0017】第2の発明は、この原理に基づく発明であり、音声情報データベースとして音声特徴量中の実音声から抽出した音響的特徴量を格納した音響情報データベースを用いることを特徴とする。

【0018】図3は本発明の第3の原理を示す原理図3であり、図中11は韻律情報検索部である。韻律情報検索部11はテキスト又は表音記号列が入力されると、実音声から抽出した様々な韻律的特徴量及びこれと対応する発声内容が格納されている韻律情報データベース21中に入力テキスト、又は入力表音記号列と一致する発声内容が存在するか否かを検索する。一致する発声内容が

存在する場合は、得られた韻律情報を音響情報生成部 31 へ渡し、また一致する発声内容が存在しない場合は、入力テキスト又は入力表音記号列をそのまま韻律情報生成部 30 へ渡す。

【0019】音響情報生成部 30 は、韻律情報生成規則 40 に基づいて入力テキスト又は入力表音記号列から韻律情報を生成し、これを音響情報生成部 31 へ渡す。音響情報生成部 31 は、音響情報生成規則 41 に基づいて韻律情報検索部 11 又は韻律情報生成部 30 から渡された韻律情報から音響情報を生成する。

【0020】第 3 の発明は、この原理に基づく発明であり、音声情報データベースとして、音声特徴量中の実音声から抽出した韻律的な特徴量を格納した韻律情報データベースを用いることを特徴とする。

【0021】図 4 は本発明の第 4 の原理を示す原理図 4 であり、図中 10 は音響情報検索部を示している。音響情報検索部 10 はテキスト又は表音記号列が入力されると実音声から抽出した様々な音響的特徴量及びこれに対応する発声内容を格納した音響情報データベース 20 中に入力テキスト又は入力表音記号列と一致する発声内容が存在するか否かを検索する。

【0022】一致する発声内容が存在する場合には直接音響情報としてこれを出力する。また一致する音声内容が存在しない場合は入力テキスト又は入力表音記号列をそのまま韻律情報検索部 11 へ渡す。韻律情報検索部 11 は、実音声から抽出した種々の韻律的特徴量及びこれに対応する発声内容が格納されている韻律情報データベース 21 中に入力テキスト又は入力表音記号列と一致する発声内容が存在するか否かを検索する。

【0023】一致する発声内容が存在する場合は、韻律的特徴量を含む韻律情報を音響情報生成部 31 へ渡し、また一致する発声内容が存在しない場合は入力テキスト又は入力表音記号列をそのまま韻律情報生成部 30 へ渡す。韻律情報生成部 30 は、韻律情報生成規則 40 に基づいて、入力テキスト又は入力表音記号列から韻律情報を生成し、これを音響情報生成部 31 へ渡す。音響情報生成部 31 は、音響情報生成規則 41 に基づいて韻律情報検索部 11 又は韻律情報生成部 30 から渡された韻律情報から音響情報を生成する。

【0024】第 4 の発明は、この原理に基づく発明であり、音声情報データベースとして、実音声から抽出した音響的特徴量を格納した音響情報データベース及び実音声から抽出した韻律的な特徴量を格納した韻律情報データベースを用いることを特徴とする。

【0025】第 5 の発明は、前記第 2、第 4 の原理に基づく発明であり、音響情報データベースとして、音声波形を格納した音声波形データベースを用いることを特徴とする。

【0026】第 6 の発明は、同じく前記第 2、第 4 の原理に基づく発明であり、音響情報データベースとして、

スペクトル、声道断面積又はフォルマント周波数を格納した音声パラメータデータベースを用いることを特徴とする。

【0027】第 7 の発明は、同じく前記第 2、第 4 の原理に基づく発明であり、音響情報データベースとして、音声波形データベース及び音声パラメータデータベースを用いることを特徴とする。

【0028】第 8 の発明は、前記第 3、第 4 の原理に基づく発明であり、韻律情報データベースとして、音声長、音声強度、基本周波数のうちの、音声長のみ、又はいずれか 2 つ以上を格納したデータベースを用いることを特徴とする。

【0029】第 9 の発明は、同じく前記第 3、第 4 の原理に基づく発明であり、韻律情報データベースとして音声長及び音声強度を格納した音声長・音声強度データベースを用いることを特徴とする。

【0030】第 10 の発明は、同じく前記第 3、第 4 の原理に基づく発明であり、韻律情報データベースとして、音声長及び基本周波数を格納した音声長・F0 データベースを用いることを特徴とする。

【0031】第 11 の発明は、同じく前記第 3、第 4 の原理に基づく発明であり、韻律情報データベースとして音声長のみを格納した音声長データベースを用いることを特徴とする。

【0032】第 12 の発明は、同じく前記第 3、第 4 の原理に基づく発明であり、韻律情報データベースとして、音声長・音声強度・F0 データベース、音声長・音声強度データベース、音声長・F0 データベース又は音声長データベースのいずれか 2 以上のデータベースを用いることを特徴とする。

【0033】

【作用】第 1 の発明にあつては、音声特徴量を格納したデータベースを用いることで高品質の合成音声の出力が可能となる。

【0034】第 2 の発明にあつては、音響的特徴量を用いることで実音声に近い明瞭度の高い合成音声を得られる。

【0035】第 3 の発明にあつては、韻律的特徴量とを用いるから、実音声に近い自然性の高い音声を得られる。

【0036】第 4 の発明にあつては、第 2、第 3 の機能を兼ね備えた機能が得られる。

【0037】第 5 の発明にあつては、音響情報データとして音声波形を用いることで自然性、明瞭度の高い音声を得られる。

【0038】第 6 の発明にあつては、明瞭度は低いが少ないデータ量で自然性の高い合成音声を得られる。

【0039】第 7 の発明にあつては、第 5、第 6 の発明の両機能を兼ねた作用が得られる。

【0040】第 8 の発明にあつては、音声長、音声強

度、F Oのうちの1又は複数を用いることで自然性の高い韻律を持った合成音声を得られる。

【0041】第9の発明にあっては、自然性の高いリズムを持った合成音声を得られる。

【0042】第10の発明にあっては、自然性の高いリズム、イントネーション、アクセントを持った合成音声を得られる。

【0043】第11の発明にあっては、少ないデータ量で自然性の高いリズムを持った合成音声を得られる。

【0044】第12の発明にあっては、自然な韻律を持った合成音声を得られる。

【0045】

【実施例】以下本発明をその実施例を示す図面にに基づき具体的に説明する。

(実施例1) 実施例1は第1、第2、第3の原理を具象化したものであり、図5は本発明の実施例1の構成を示すブロック図である。図5中100は音声波形検索部を示している。

【0046】音声波形検索部100はテキスト又は表音記号列が入力されると、音声波形データベース200を検索する。音声波形データベース200には発声内容を示す表音記号列と、それに対応した音声波形データ（PCMデータ又はADPCM等の符号化技術により情報圧縮されたデータ）の対が複数格納されて、この内に入力テキスト又は入力表音記号列と一致する発声内容が存在する場合には、対応する音声波形を直ちにDA変換部5へ渡し、また一致する発声内容が存在しない場合には入力テキスト又は入力表音記号列をそのまま音声パラメータ検索部101へ渡す。

【0047】音声パラメータ検索部101は、実音声から抽出した音声パラメータ、例えばPARCOR、LSP、フォルマント周波数、及びこれに対応する発声内容を格納した音声パラメータデータベース201中に入力テキスト又は入力表音記号列と一致する発声内容が存在するか否かを検索する。

【0048】一致する発声内容が存在すれば、その音声パラメータを波形生成部311へ渡し、また一致する発声内容が存在しない場合は入力テキスト又は入力表音記号列をそのまま韻律情報検索部11へ渡す。韻律情報検索部11は、実音声から抽出した種々の韻律的特徴量及びこれに対応する発声内容を格納した韻律情報データベース21中に入力テキスト又は表音記号列と一致する発声内容が存在するか否かを検索する。一致する発声内容が存在すれば、その韻律情報を音声パラメータパターン生成部310へ渡し、また一致する発声内容が存在しない場合は入力テキスト又は入力表音記号列をそのまま韻律情報生成部30へ渡す。

【0049】韻律情報生成部30は韻律情報生成規則40に基づいて、入力テキスト又は入力表音記号列から韻律情報を生成し、これを音声パラメータパターン生成部

310へ渡す。音声パラメータパターン生成部310は音声パラメータ生成規則410に基づいて、渡された韻律情報、即ち表音記号列（特に読み情報）、音声長及び音声強度パターン、F Oパターンから音声パラメータパターンを生成し、これを波形生成部311へ渡す。生成される音声パラメータパターンは具体的にはRARCOR係数、又はフォルマント周波数と音源信号との時間変化パターンであったり、所謂波形編集方式の場合にあっては、音声波形の短い単位である素片波形の種類、接続タイミング等の接続情報である。

【0050】波形生成部311は、音声パラメータ検索部101又は音声パラメータ生成部310から渡された音声パラメータパターンに基づいて実際の音声波形を生成し、DA変換部5へ渡す。音声波形の生成処理は、音声パラメータパターンがRARCOR係数の場合は波形生成部311はPARCORフィルタと音源生成部とからなり、RARCORフィルタを音源信号で駆動処理し、また波形編集方式の場合は素片波形を適切な位置に配し、これらを滑らかに接続する処理を行う。DA変換部5は波形生成部311で生成され、又は音声波形検索部100によって音声波形データベース200から引き出されたデジタル音声波形をアナログ音声波形に変換して合成音声として出力する。

【0051】このような実施例1にあっては、音声波形データベース200及び音声パラメータデータベース201を用いることによって高品質な合成音声の生成が、また韻律情報データベース21を用いることで自然な合成音声の生成が可能となる。

【0052】（実施例2）実施例2は第3の原理を具象化したものであり、図6は、本発明の実施例2の構成を示すブロック図である。図6中110は音声長・音声強度・F Oパターン検索部を示している。音声長・音声強度・F Oパターン検索部61はテキスト又は表音記号列が入力されると、実音声から抽出した音声長・音声強度・F Oパターン及びこれに対応する発声内容を格納した音声長・音声強度・F Oデータベース210中に、入力テキスト又は入力表音記号列と一致する発声内容が存在するか否かを検索する。

【0053】一致する発声内容が存在すれば、その音声長・音声強度・F Oパターンを音響情報生成部31へ渡し、また一致する発声内容が存在しない場合には入力テキスト又は入力表音記号列をそのまま音声長・F Oパターン検索部111へ渡す。音声長・音声強度・F Oデータベースには、発声内容を示す情報とそれに対応した音素及び音節等の合成単位毎の時間長である音声長パターンと、音声強度の時間変化パターンである音声強度パターンと、基準周波数の時間変化パターンであるF Oパターンとの組が複数格納されている。

【0054】音声長・F Oパターン検索部111は音声長・F Oデータベース211中に入力テキスト又は入力

表音記号列と一致する発声内容が存在するか否かを検索する。一致する発声内容が存在する場合は、その音声長・F Oパターンデータを音声強度パターン生成部302に渡し、また一致する発声内容が存在しない場合は、入力テキスト又は入力表音記号列をそのまま音声長・音声強度パターン検索部112へ渡す。音声長・音声強度パターン検索部112は音声長・音声強度データベース212中に入力テキスト又は入力表音記号列と一致する発声内容が存在するか否かを検索し、一致する発声内容が存在する場合はその音声長・音声強度パターンをF Oパターン生成部301へ渡し、また存在しない場合は入力テキスト又は入力表音記号列をそのまま音声長パターン検索部113へ渡す。

【0055】音声長パターン検索部113は、音素、音節等の合成単位毎の時間長である音声長パターンとこれに対応する発声内容が格納された音声長データベース213中に、入力テキスト又は入力表音記号列が存在するか否かを検索し、存在する場合はこれをF Oパターン生成部301へ渡し、また存在しない場合は入力テキスト又は入力表音記号列をそのまま音声長パターン生成部300へ渡す。音声長パターン生成部300は音声長生成規則400に基づいて入力テキスト又は入力表音記号列を音声長パターンに変換し、これをF Oパターン生成部301へ渡す。

【0056】F Oパターン生成部301はF O生成規則401に基づいて、音声長・音声強度パターン、又は音声長パターンと表音記号列とからアクセント、イントネーションに相当する物理量であるF Oパターンを生成し、表音記号列と共にこれを音声強度パターン生成部302へ渡す。音声強度パターン生成部302は音声強度生成規則402に基づいて、表音記号列、音声長・F Oパターン又はF Oパターンから音声強度の時間変化パターンである音声強度パターンを生成し、これを音響情報生成部31へ渡す。音響情報生成部31は音響情報生成規則41に基づいて実際の音響情報を生成する。

【0057】このような実施例2にあつては、音声長・音声強度・F Oデータベース210、音声長・F Oデータベース211、音声長・音声強度データベース212及び音声長データベース213等を用いることで、自然なイントネーション、アクセント及びリズムを持った合成音声の生成が可能となる。

【0058】(実施例3) 実施例3は第1、第2、第3、第4の原理を具象化したものであり、図7は実施例3の構成を示すブロック図である。図7中304は言語処理部を示している。言語処理部304はテキスト入力があると単語辞書303を参照して入力テキスト中の単語の読み、アクセント情報及び文の構成を解析し、イントネーションを制御するための制御情報からなる表音記号列を作成し、これを音声波形検索部100へ渡す。

【0059】音声波形検索部100は言語処理部で作成

された表音記号列をキーとして音声波形データベース200からキーと一致する発声内容に対応する音声波形データを検索し、一致する発声内容が存在する場合はその音声波形データをD A変換部5に渡し、また存在しない場合には入力表音記号列をそのまま音声パラメータ検索部101へ渡す。音声パラメータ検索部101は入力表音記号列をキーとして、音声パラメータデータベース201からキーと一致する発声内容に対応する音声パラメータデータを検索し、一致する発声内容が存在する場合はその音声パラメータデータを波形生成部311へ渡し、また一致しない場合は入力表音記号列をそのまま音声長・F Oパターン検索部110へ渡す。

【0060】音声長・F Oパターン検索部110は入力表音記号列をキーとして、音声長、F Oデータベース210からこれと一致する発声内容に対応する音声長、F Oデータを検索し、一致する発声内容が存在すれば、その音声長・F Oデータを音声パラメータパターン生成部310へ渡し、また存在しない場合には入力された表音記号列をそのまま音声長パターン検索部111へ渡す。

【0061】音声長・音声強度・F Oデータベース210には、音声内容を示す表音記号列とそれに対応した音素、音節等の合成単位毎の時間長である音声長パターンと、基本周波数の時間変化パターンであるF Oパターンとの対が複数格納されている。音声長パターン検索部111は入力表音文字列をキーとして、音声長データベース211からキーと一致する発声内容に対応する音声長データを検索し、一致する発声内容が存在する場合は、それをF Oパターン生成部301へ渡し、また存在しない場合には入力表音記号列をそのまま音声長パターン生成部300へ渡す。音声長パターン生成部300は音声長生成規則400に基づいて入力表音記号列を音声長パターンに変換し、これを入力表音記号列と共にF Oパターン生成部301へ渡す。

【0062】F Oパターン生成部301はF O生成規則401に基づいて入力表音記号列と音声長パターンとから、アクセント、イントネーションに相当する物理量であるF Oパターンを生成し、これを入力表音記号列と共に音声パラメータパターン生成部310へ渡す。音声パラメータパターン生成部310は音声パラメータ生成規則410に基づいて入力表音記号列と音声長・F Oパターン又はF Oパターンから音声パラメータパターンを生成し、これを波形生成部311へ渡す。

【0063】波形生成部311は、音声パラメータ検索部101又は音声パラメータパターン生成部310からされた音声パラメータパターンからデジタルに音声波形を生成し、D A変換部5へ渡す。D A変換部5は波形生成部311で生成され、又は音声波形検索部100にて検索されたデジタル音声波形をアナログ音声波形に変換し、合成音声として出力する。

【0064】このような実施例3にあつては、最初に言

語処理部 304 による言語処理を行い、入力テキストを読み、アクセントを示す表音記号列に変換した後、音声波形データベースの検索を行うから同じ単語が漢字、平仮名、片仮名、異なる送りがな等種々の表記で書かれていてもデータベースの容量削減が出来る効果がある。

【0065】（実施例 4）実施例 4 は第 1、第 2、第 3、第 4 の原理を具象化したものであり、図 8 は実施例 4 の構成を示すブロック図である。この実施例 4 にあっては、図 7 に示す実施例 3 において、言語処理部 304 を音声長パターン検索部 111 と音声長パターン生成部 300 との中間に位置させたのと実質的に同じ構成となっている。このような実施例 4 については、いま「こんにちは。」「今日は 3 月 3 日です。」「ありがとうございました。」の 3 文が入力された場合について具体的にその処理過程を説明する。

【0066】なお「こんにちは」の音声波形データは音声波形データベース 200 に格納され、「ありがとうございました」の音声長・FO パターンは音声長・FO データベース 210 に格納されているものとし、また「今日は 3 月 3 日です。」はいずれのデータベースにも格納されていないものとする。いま上述した 3 文が入力された場合、「こんにちは」の文が入力されると音声波形検索部 100 が音声波形データベース 200 を検索し、ここに格納されている音声波形データを検索し、この音声波形データを直接 DA 変換部 5 へ送り、合成音声「コンニチハ」として出力される。

【0067】また「ありがとうございました」の場合には、音声波形検索部 100、音声パラメータ検索部 101 では検索されないが、音声長・FO パターン検索部 110 が音声長・FO データベース 210 を検索することでこれを検出し、これを音声パラメータパターン生成部 310 へ渡し、波形生成部 311 を経て合成音声として「アリガトウゴザイマシタ」が出力される。一方「今日は 3 月 3 日です」の場合はいずれのデータベースにも格納されていないから言語処理部 304 に達し、ここで単語の読み、アクセント位置を決定しその文の構造を解析し、イントネーションを制御する制御情報を音声パターン生成部 300 へ渡す。音声情報パターン生成部 300 は音声長生成規則 400 に基づき音声長パターンを生成し、これを FO パターン生成部 301 へ渡す。

【0068】以下実施例 3 と同様にして、FO パラメータパターン生成部 301、音声パラメータパターン生成部 310、波形生成部 311 を経て DA 変換部 5 へ渡され、「キョウハサンガツミツカデス」の合成音声出力される。

【0069】このような実施例 4 にあっては、最初に音声波形データベース 200、音声パラメータデータベース 201、音声長・FO データベース 210、音声長データベース 211 の検索を行うことで、入力テキストがデータベース内に存在する場合、それ以降の検索処理を

行う必要がなく、高速に合成音声の生成が可能となる。

【0070】（実施例 5）実施例 5 は第 1、第 2、第 3、第 4 の原理を具象化したものであり、図 9 は、実施例 5 の構成を示すブロック図である。図 9 中 10a、10b は音響情報検索部を、11a、11b は韻律情報検索部を夫々示している。音響情報検索部 10a は、テキスト入力されると入力テキストをキーにして音響情報データベース 20 を検索し、入力テキストと一致する発声内容が存在する場合は、これを音響情報として直接出力する。また、一致する発声内容が存在しない場合は、そのまま入力テキストを韻律情報検索部 11a へ渡す。

【0071】韻律情報処理 11a は入力テキストをキーにして韻律情報データベース 21 を検索し、入力テキストと一致する発声内容が存在する場合はそれを音響情報生成部 31 へ渡し、また一致する発声内容が存在しない場合はそのまま言語処理部 304 へ渡す。言語処理部 304 は単語辞書 303 を参照しつつ、入力テキストを解析し、表音記号列に変換し、これを音響情報検索部 10b へ渡す。

【0072】音響情報検索部 10b は表音記号列をキーにして音響情報データベース 20 を検索し、表音記号列と一致する発声内容が存在する場合は、韻律情報データベースを検索の生成、韻律情報の生成は行うことなく、音響情報データベース 20 から取り出した音響情報を直接出力する。また、一致する発声内容が存在しない場合は、そのまま音響情報を韻律情報検索部 11b へ渡す。

【0073】韻律情報検索部 11b は表音記号列をキーにして韻律情報データベース 21 を検索し、表音記号列と一致する発声内容が存在する場合はこれを直に音響情報として音響情報生成部 30 へ渡し、また存在しない場合はそのまま韻律情報生成部 30 へ渡す。韻律情報生成部 30 は、韻律情報生成規則 40 に基づいて入力表音記号列から韻律情報を生成し、これを音響情報生成部 31 へ渡す。韻律情報生成部 31 は音響情報生成規則 41 を用いて韻律情報から音響情報を生成し、これを出力する。

【0074】このような実施例 5 にあっては、最初に音響情報データベース 20、韻律情報データベース 21 の検索を行うから、入力テキストが音響情報データベース 20、又は韻律情報データベース 21 内に存在する場合、それ以降の処理を省略出来ると共に、データベースの検索後に言語処理を行い、入力テキストを表音記号列に変換した後、音響情報データベース、韻律情報データベースの検索を行うから、同じ単語が種々異なる表記で書かれていてもデータベースに格納された発声内容とのマッチングが可能となる。

【0075】

【発明の効果】第 1 の発明にあっては、実音声进行分析して抽出した音声特徴量を格納した音声情報データベースを用いることで、高品質な合成音声の出力が可能となる

ことは勿論、データベースに存在しない場合には規則による音声合成を行うことで広範囲の入力テキスト、入力表音記号列からの合成音声の生成が可能となる。

【0076】第2の発明にあつては、実音声进行分析し、抽出した音響的特徴量を用いることで、実音声に近い明瞭度の高い合成音声を得られる。

【0077】第3の発明にあつては、実音声进行分析して抽出した韻律的特徴量を用いることで、実音声に近い自然性の高い合成音声を得られる。

【0078】第4の発明にあつては、第2、第3の発明の両効果を備える。

【0079】第5の発明にあつては、音響データベースとして音声波形を用いることで、自然性と明瞭性の高い合成音声を得られる。

【0080】第6の発明にあつては、第5の発明よりも明瞭度は低い、少ないデータ量で明瞭度の高い合成音声を得られる。

【0081】第7の発明にあつては、第5、第6の発明の効果を兼ね備え得る。

【0082】第8の発明にあつては、韻律情報として音声長・音声強度・F0を用いることで、自然性の高い韻律情報を持った合成音声を得られる。

【0083】第9の発明にあつては、韻律情報として音声長・音声強度を用いることで、自然性の高いリズムを持った合成音声を得られる。

【0084】第10の発明にあつては、韻律情報データとして音声長・F0パターンを用いることで、自然性の高いリズム、イントネーション、アクセントを合成音声を得られる。

【0085】第11の発明にあつては、韻律情報データとして音声長を用いることで、少ないデータ量で自然性の高いリズムを持った合成音声を得られる。

【0086】第12の発明にあつては、より自然な韻律を持った合成音声を得られる。

【図面の簡単な説明】

【図1】本発明の原理を示す原理図である。

【図2】本発明の他の原理を示す原理図である。

【図3】本発明の更に他の原理を示す原理図である。

【図4】本発明の更に他の原理を示す原理図である。

【図5】本発明の実施例1の構成を示すブロック図である。

【図6】本発明の実施例2の構成を示すブロック図である。

【図7】本発明の実施例3の構成を示すブロック図である。

【図8】本発明の実施例4の構成を示すブロック図である。

【図9】本発明の実施例5の構成を示すブロック図である。

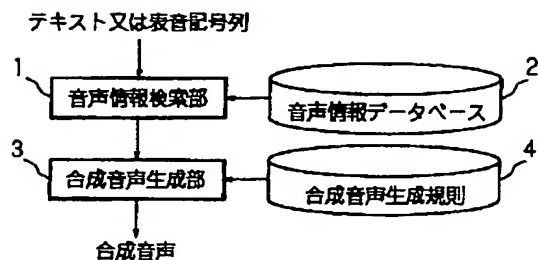
【図10】従来の音声合成システムの構成を示すブロック図である。

【符号の説明】

- | | |
|-----|--------------------|
| 1 | 音声情報検索部 |
| 2 | 音声情報データベース |
| 3 | 合成音声生成部 |
| 4 | 合成音声生成規則 |
| 5 | DA変換部 |
| 10 | 音声情報検索部 |
| 11 | 韻律情報検索部 |
| 20 | 音声情報データベース |
| 21 | 韻律情報データベース |
| 30 | 韻律情報生成部 |
| 31 | 音響情報生成部 |
| 40 | 韻律情報生成規則 |
| 41 | 音響情報生成規則 |
| 100 | 音声波形検索部 |
| 101 | 音声パラメータ検索部 |
| 110 | 音声長・音声強度・F0パターン検索部 |
| 111 | 音声長・F0パターン検索部 |
| 112 | 音声長・音声強度パターン検索部 |
| 113 | 音声長パターン検索部 |
| 200 | 音声波形データベース |
| 201 | 音声パラメータデータベース |
| 210 | 音声長・音声強度・F0データベース |
| 301 | F0パターン生成部 |
| 302 | 音声強度パターン生成部 |
| 310 | 音声パラメータパターン生成部 |
| 311 | 波形生成部 |

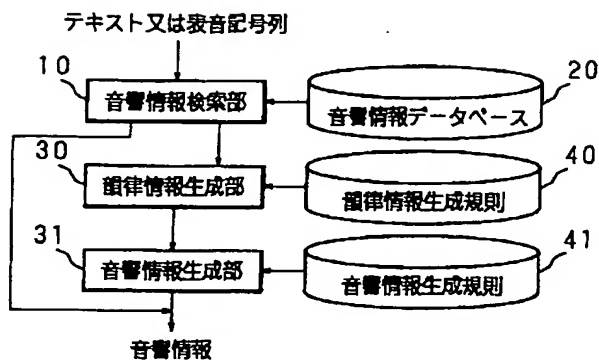
【図 1】

本発明の原理図 1



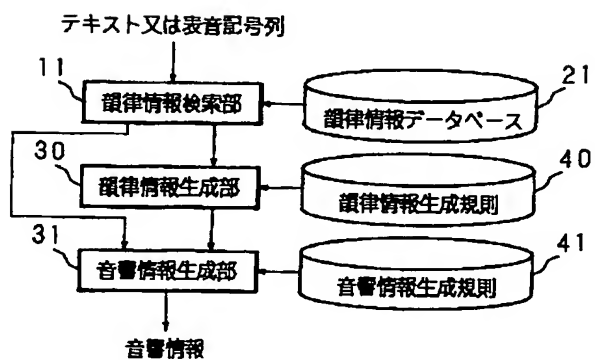
【図 2】

本発明の原理図 2



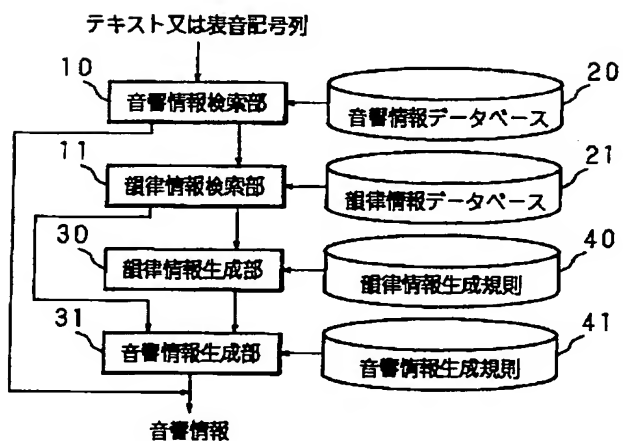
【図 3】

本発明の原理図 3



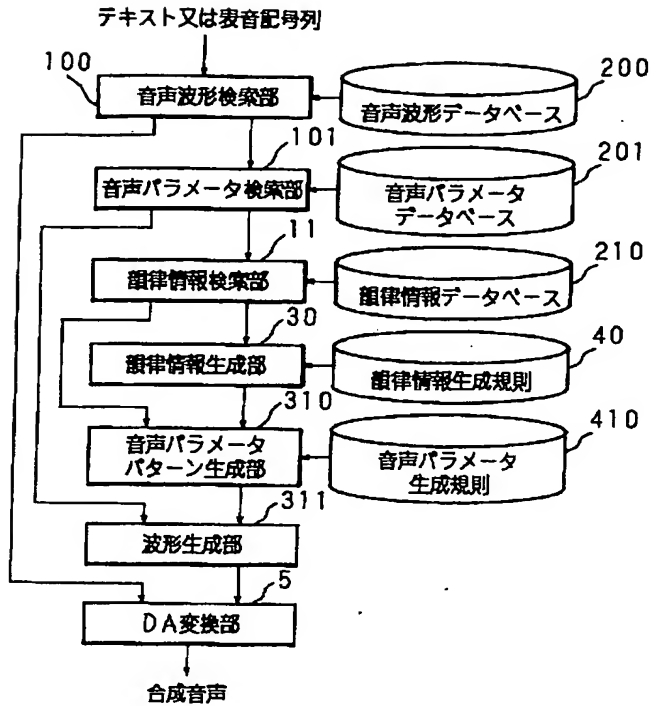
【図 4】

本発明の原理図 4



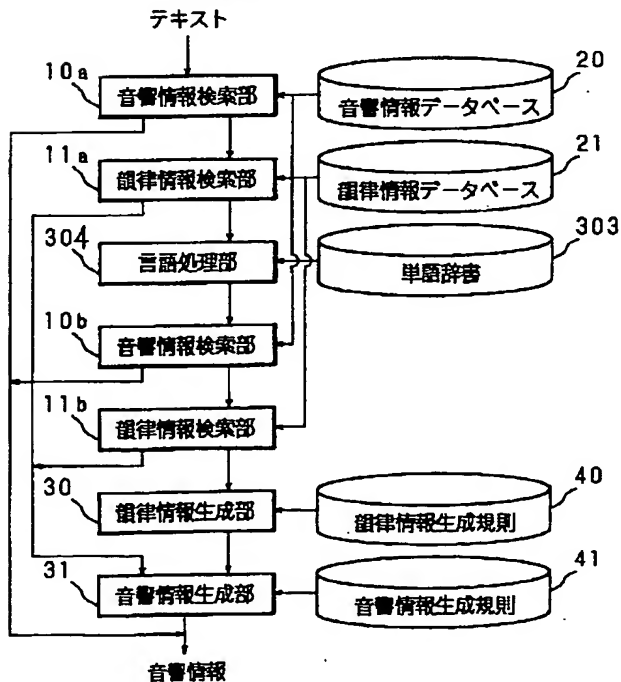
【図 5】

実施例 1 のブロック図



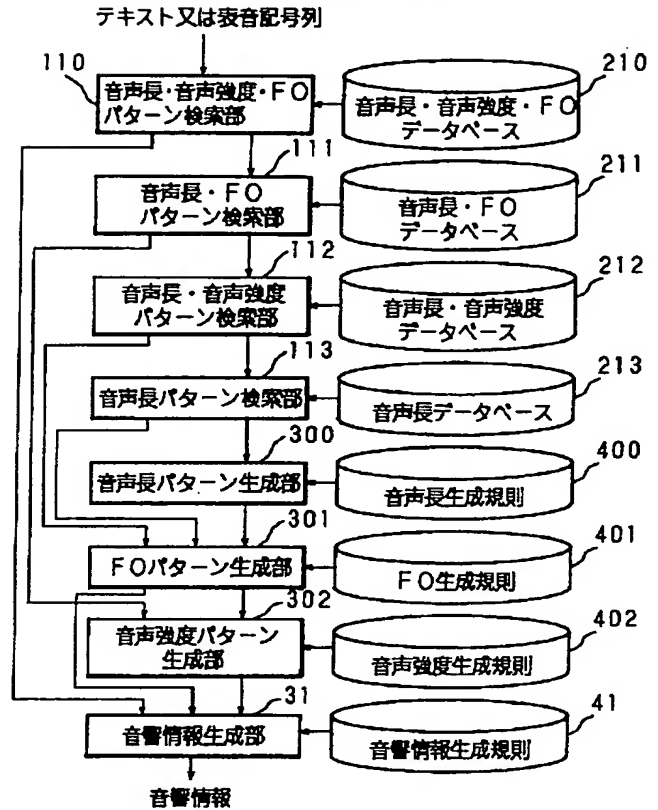
【図 9】

実施例 5 のブロック図



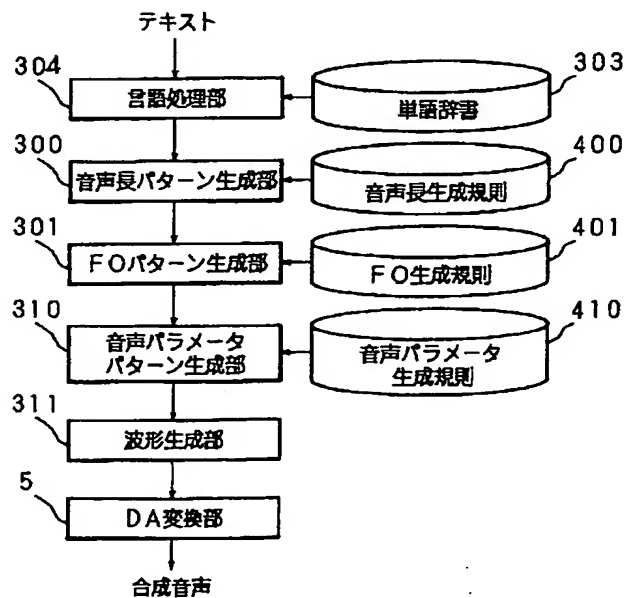
【図 6】

実施例 2 のブロック図



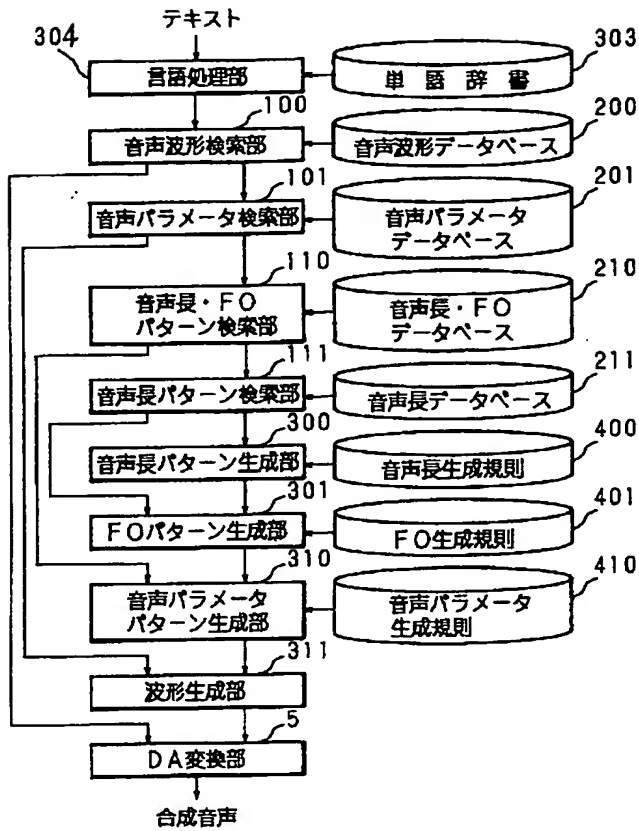
【図 10】

従来例 1 のブロック図



【図 7】

実施例 3 のブロック図



【図 8】

実施例 4 のブロック図

